

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-6706

(43)公開日 平成9年(1997)1月10日

(51) IntCl'

G O 6 F 13/14
13/00

識別記号

3 1 0
3 5 5

序内整理番号

7368-5E
7368-5E

FI

G O B F 13/14
13/00

技術表示箇所

310Y
355

審査請求 未請求 請求項の数22 O.L (全 18 頁)

(21)出願番号

特開平7-155762

(22) 出題日

平成7年(1995)6月22日

(71) 出題人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 村山 秀樹

東京都国分寺市東盛ヶ丘1丁目280番地
株式会社日立製作所中央研究所内

(72) 発明者 屋代 寛

東京都国分寺市東恋ヶ窪1丁目280番地
株式会社日立製作所中央研究所内

(72) 究明者 吉沢 聡

東京都国分寺市東恋ヶ窪1丁目280番地
株式会社日立製作所中央研究所内

(74) 代理人 弁理士 小川 勝男

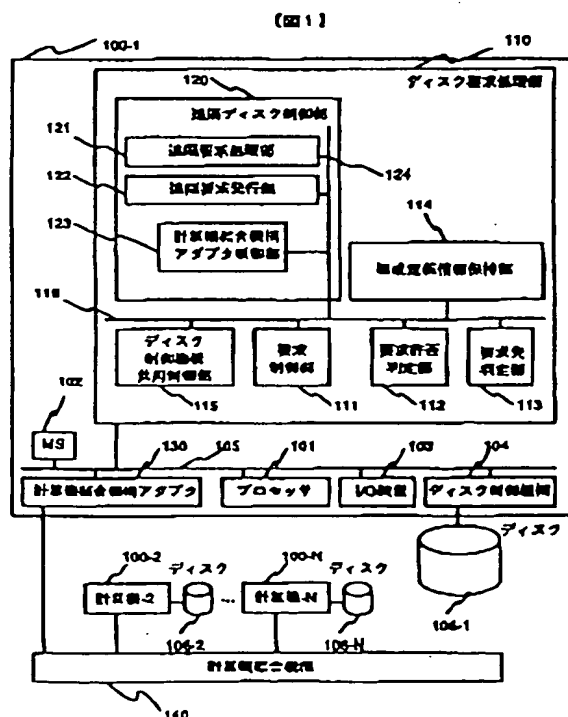
最終頁に読む

(54)【発明の名称】 疎結合計算機システム

(57) 【要約】 (修正有)

【目的】 プロセッサが接続機構を介して共用ハードディスクを直接制御できる。

【構成】 疎結合計算機システムは、計算機１００－１～Ｎが、計算機結合機構１４０によって相互に結合されて構成される（Ｎは自然数とする以下同様）。計算機１００－１～Ｎは、プロセッサ１０１、メモリ１０２、Ｉ／Ｏ装置１０３、ディスク制御機構１０４、計算機結合網アダプタ１３０、ディスク要求処理部１１０を含む。ディスク要求処理部１１０は、必要に応じて計算機アダプタ１０５及び計算機結合機構１４０を介してディスク要求処理部１１０間の通信を行い、プロセッサ１０１からの処理要求を対応するディスク制御機構１０４に要求する。ディスク要求処理部１１０がディスク制御機構１０４に対する処理要求を代行して要求することにより、プロセッサ１０１は、すべてのディスク１０６－１～Ｎ）に対するアクセスを同一のアクセス制御手順にて制御可能となる。



(2)

特開平9-6706

1

【特許請求の範囲】

【請求項1】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を制御するために、他の計算機からの制御要求及び他の計算機への制御要求がある場合に上記計算機結合機構を介して接続することにより複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項2】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を制御するために、他の計算機からの制御要求及び他の計算機への制御要求がある場合に上記計算機結合機構を介して接続し、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に制御を行うことにより複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項3】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持することで、上位ソフトウェア及びハードウェアの接続を変更することなく、動作する実装置を変更することが可能な疎結合計算機システムで、処理が必要な場合上記計算機結合機構を介して接続し、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に制御を行うことにより複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項4】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、上記計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項5】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて上記計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項6】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、上記

2

計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項7】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持しており、上記計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項8】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで上記計算機結合機構を制御する切り替え部分の制御情報を変更することにより、動作する実装置の変更による処理を増加することなく、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項9】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで上記計算機結合機構を制御する切り替え部分の制御情報を変更することにより、動作する実装置の変更による処理を増加することなく、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項10】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで上記計算機結合機構を制御する切り替え部分の制御情報を変更することにより、上位ソフトウェアを変更することなく、対応する実装置の変更による処理を増加することなく、複数の

(3)

特開平9-6706

3

計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項11】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで上記計算機結合機構を制御する切り替え部分の制御情報を変更することにより、ハードウェアの接続を変更することなく、動作する実装置の変更による処理を増加することなく、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項12】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで上記計算機結合機構を制御する切り替え部分の制御情報を変更することにより、上位ソフトウェア及びハードウェアの接続を変更することなく、動作する実装置の変更による処理を増加することなく、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項13】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで、対応する実装置を変更することが可能な、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項14】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで、ハードウェアの接続を変更することなく、対応する実装置を変更することが可能な、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項15】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで、上位ソフトウェアを変更することなく、対応する実装置を変更することが可能な、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

4

【請求項16】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで、上位ソフトウェア及びハードウェアの接続を変更することなく、対応する実装置を変更することが可能な、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項17】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、この構成定義情報を変更することで、上位ソフトウェア及びハードウェアの接続を変更することなく、対応する実装置を変更することが可能な、構成定義情報に保持された制御権限を調べて制御の可否を判定する要求許否判定部を持ち、制御権限があることが確認された場合に複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項18】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、装置共用を行うことを可能とすることを特徴とする疎結合計算機システム。

【請求項19】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持し、複数の計算機による装置共用を行うことを特徴とする疎結合計算機システム。

【請求項20】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持することで、上位ソフトウェア及びハードウェアの接続を変更することなく、対応する実装置を変更することが可能で、上記計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項21】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮

50

(4)

特開平9-6706

5

想的な識別子と実装置との対応を構成定義情報に保持して、この構成定義情報を変更することで切り替え部分の制御情報を変更することにより、上位ソフトウェア及びハードウェアの接続を変更することなく、対応する実装置を変更することが可能で、上記計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【請求項22】複数の計算機が計算機結合機構を用いて相互に疎結合された疎結合計算機システムにおいて、各計算機に各々接続された共用二次記憶装置を全ての計算機から同一に扱える仮想的な識別子を用いて管理し、仮想的な識別子と実装置との対応を構成定義情報に保持して、この構成定義情報を変更することで切り替え部分の制御情報を変更することにより、上位ソフトウェア及びハードウェアの接続を変更することなく、対応する実装置を変更することが可能で、処理が必要な場合に上記計算機結合機構を介して接続することで可能で、計算機結合機構を介して相互に保持している装置に対して遠隔処理要求を発行する部分と遠隔処理要求を処理する部分を備え、各計算機が相互に他の装置を制御することを特徴とする疎結合計算機システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、計算機システムを相互に結合して構成された疎結合計算機システムに関する。

【0002】

【従来の技術】疎結合計算機システムでは、データを共有するためにハードディスクを共有するシステムが提案されている。ハードディスクを共有するためには、複数のハードディスク制御装置と共有対象となるハードディスクを接続する接続機構が必要である。例えば、文献「ナンシー・ビー・クロネンバーグ他、ザ・バックスクラスタ概念：アン・オーバービュー・オブ・ディストリビューテッド・システム、7～21頁、デジタル・テクニカル・ジャーナル、No. 5、1987年9月、デジタル・イクイップメント・コーポレーション」(Nancy P. Kronenberg et. al, The VAXcluster Concept: An Overview of a Distributed System, pp7-21, Digital Technical Journal, No. 5 September 1987 Digital Equipment Corporation)やANSIにて標準化が進んでいるFibre ChannelのSCSI (Small Computer System Interface) Profileによるハードディスク共有が文献「ロバート・スニブリー、インプリメンティング・ファイバーチャネ

6

ル・スカジ・トランスポート、78～82頁、コンプコン、春94」(Robert Snively, Implementing a Fibre Channel SCSI Transport, pp78-82, COMPCON, Spring 94)がある。

【0003】また、疎結合計算機システムにおいて、RPC (Remote Procedure Call) によりファイル共用を実現する手段としてNFS (Network File System) 「ジョージ・エフ・コーロリス他、分散システムコンセプトとデザイン、391～407頁」(George F. Coulouris et. al, DISTRIBUTED SYSTEMS, pp391-407, 1991)などがある。

【0004】

【発明が解決しようとする課題】Fibre Channelでは、共有ハードディスクにアクセスする際には必ず接続機構を介する必要があるため、接続機構に対するアクセスが増加することにより性能低下を招くという問題があった。

【0005】また、NFS (Network File System) は、他のプロセッサに処理を代行させる方式であり、処理の代行要求を行うためにRPCによるプロセッサ間の通信を行うためにソフトオーバーヘッドが増加し、性能が低いという問題があった。

【0006】本発明の目的は、プロセッサが接続機構を介して共用ハードディスクを直接制御でき、高性能なハードディスク共有が可能な疎結合計算機システムを構成することである。

【0007】

【課題を解決するための手段】本発明では、ソフトウェアからの要求先となるハードディスクを判定する要求先判定部を持ち、対応するハードディスクに対する処理を要求するハードディスク要求処理部を持ち、ハードディスクの接続するバスとリモートアクセスのための接続機構の間の処理要求を必要に応じて転送する遠隔ディスク制御部を持ち、ハードディスクアクセス要求があるプロセッサがハードディスクを直接制御することができる構成とする。

【0008】

【作用】本発明によれば、共有されるハードディスクへのアクセスであっても、特定の計算機からのアクセスについては、接続機構に対するアクセスが発生せず、要求の発生したプロセッサがハードディスクを直接制御できる。

【0009】

【実施例】以下図を用いて本発明の実施例を説明する。

【0010】図1に実施例の一つを示す。疎結合計算機システムは、計算機(100-1から100-N)が、計算機結合機構140によって相互に結合されて構成さ

(5)

特開平9-6706

7

れる(Nは自然数とする以下同様)。計算機(100-1から100-N)は、プロセッサ101、メモリ102、I/O装置103、ディスク制御機構104、計算機結合網アダプタ130、ディスク要求処理部110がシステムバス105に結合し、ディスク(106-1から106-N)がディスク制御機構104に接続して構成される。プロセッサ101は、例えばCPUによって構成できる。メモリ102は、例えばRAMによって構成できる。I/O装置103は、入出力をするための装置であり、例えばディスプレイ装置、キーボード、ハードディスク等の二次記憶によって構成される。I/O装置103は、計算機結合網アダプタ130による通信処理によって入出力機能を代替する構成の場合は、計算機(100-1から100-N)の要素として含まれない場合もある。

【0011】ディスク制御機構104は、ディスク要求処理部110からの処理要求に対してディスク(106-1から106-N)を制御する機構であり、たとえば、TTL、CMOS等の論理によって構成できる。ディスク(106-1から106-N)は、ハードディスク等の二次記憶によって構成できる。

【0012】計算機結合網アダプタ130は、計算機(100-1から100-N)と計算機結合機構140とを相互に結合し、プロセッサ101及びディスク要求処理部110の指示によりデータの送受信処理を行なう部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0013】計算機結合機構140は、計算機(100-1から100-N)を結合して相互通信をするための結合網であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0014】ディスク要求処理部110は、必要に応じて計算機アダプタ105及び計算機結合機構140を介してディスク要求処理部110間の通信を行い、プロセッサ101からの処理要求を対応するディスク制御機構104に要求する。ディスク要求処理部110がディスク制御機構104に対する処理要求を代行して要求することにより、プロセッサ101は、すべてのディスク(106-1から106-N)に対するアクセスを同一のアクセス制御手順にて制御可能となる。また、共有が必要な場合のみディスク(106-1から106-N)の共有ができるために、計算機結合機構140が性能ネックになることを避けることができる。

【0015】次に、ディスク要求処理部110の詳細を説明する。ディスク要求処理部110は、同一計算機内のプロセッサ101による処理要求と、計算機結合機構140を介して接続されている計算機(100-1から100-N)からの遠隔処理要求との2種類の処理要求を受け、同一計算機内のディスク制御機構104に対する制御と、計算機結合機構140を介して接続されてい

8

る計算機(100-1から100-N)に対する遠隔制御要求の発行の2種類の制御処理を行う。ディスク要求処理部110は、要求制御部111、要求許否判定部112、要求先判定部113、構成定義情報保持部114、ディスク制御機構共用制御部115、リモートディスク制御部120が内部バス116に接続して構成される。

【0016】要求制御部111は、ディスク制御機構104に対する同一計算機内のプロセッサ101による処理要求と、計算機結合機構140を介して接続されている計算機からの遠隔処理要求を受け付けて、制御を行う部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0017】要求許否判定部112は、構成定義情報保持部114に保持された構成定義情報に基づいて、ディスク要求処理部110に対して要求された処理の許否を判定する処理を行う部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0018】要求先判定部113は、構成定義情報保持部114に保持された構成定義情報に基づいて、ディスク要求処理部110に対して要求された処理をどの計算機のどのディスクに対する処理であるかを判定する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0019】構成定義情報保持部114は、ディスク要求処理部110に対して要求された処理がどの計算機のどのディスクに対応するか、処理の要求者が対応するディスクに対するアクセスの権利(アクセス権)を持っているか、セキュリティを向上させるためのパスワード等の構成定義情報を保持する部分であり、例えばRAMや電気的に書換可能なROM等によって構成できる。

【0020】ディスク制御機構共用制御部115は、ディスク制御機構104に対する同一計算機内のプロセッサ101による処理要求と、計算機結合機構140を介して接続されている計算機からの遠隔処理要求との2種類の処理要求の間の調停を行う部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0021】遠隔ディスク制御部120は、計算機結合機構140を介して接続されている計算機からの遠隔処理要求に対するディスク制御機構104の制御と計算機結合機構140を介して接続されている計算機に対する遠隔制御要求の発行を行う部分である。

【0022】遠隔ディスク制御部120は、遠隔要求処理部121、遠隔要求発行部122、計算機結合機構アダプタ制御部123が内部バス124に接続して構成される。遠隔要求処理部121は、計算機結合機構140を介して接続されている計算機からの遠隔処理要求に対する処理を行う部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0023】遠隔要求発行部122は、計算機結合機構

(6)

特開平9-6706

9

10

140を介して接続されている計算機に対して遠隔制御要求を発行する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0024】計算機結合機構アダプタ制御部123は、計算機結合機構140を介して接続されている計算機に対する通信を行うために、計算機結合機構アダプタ105を制御する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0025】図2に疎結合計算機システム構成の他の例2を示す。図2は、三台の計算機とおのおのに接続されたディスクが部分的に共有された場合の例である。計算機-1(100-1)は、ディスク(106-1(ID=A, B, C)) : IDは、各計算機内部における識別子を有し、ディスク(106-1(ID=A, B))は非共有ディスクであり、ディスク(106-1(ID=C))が計算機-1(100-1)と計算機-2(100-2)によって共有されるディスクである。

【0026】計算機-2(100-2)は、ディスク(106-2(ID=A, B))を有し、ディスク(106-2(ID=A))は非共有ディスクであり、ディスク(106-2(ID=B))が計算機-2(100-2)と計算機-3(100-3)によって共有されるデ*

表1

| 仮想ディスク | 実ディスク管理計算機 | 管理計算機内対応ディスク | アクセス権情報 | パスワード |
|---------|------------|--------------|---------------------|-------|
| VDISK-1 | 計算機-1 | A | 計算機-1 | ABC |
| VDISK-2 | 計算機-2 | A | 計算機-2 | BCD |
| VDISK-3 | 計算機-3 | A | 計算機-3 | CDE |
| VDISK-4 | 計算機-3 | C | 計算機-1, 計算機-2, 計算機-3 | DEF |
| VDISK-5 | 計算機-3 | B | 計算機-1, 計算機-2, 計算機-3 | EFG |
| VDISK-6 | 計算機-2 | B | 計算機-2, 計算機-3 | FGH |
| VDISK-7 | 計算機-1 | B | 計算機-1 | GHI |
| VDISK-8 | 計算機-1 | C | 計算機-1, 計算機-2 | HIJ |

【0030】実ディスク管理計算機とは、実際に処理コマンドが発行される対象となるディスク(106-1から106-3)を制御するディスク制御機構104が接続されている計算機(100-1から100-3)を示す。

【0031】管理計算機内対応ディスクとは、仮想ディスクに対応する実ディスク管理計算機内のディスク(106-1から106-3)を示している。以下、これを実ディスクという。この例では、管理計算機がディスクを管理するための識別子(ID)の形式で格納されていることを示す。アクセス権情報は、各仮想ディスクに対して、どの計算機からのアクセスを許可するかを示す。この例では、アクセスを許可する計算機の識別子である。アクセス権情報は、データ読み込み、データ書き込み、状態読み出し等の各処理コマンドごとに保持されている。パスワードは、セキュリティを向上するために使用する。たとえば、信頼性を向上させるために、※50

*リスクである。

【0027】計算機-3(100-2)は、ディスク(106-3(ID=A, B, C))を有し、ディスク(106-3(ID=A))は非共有ディスクであり、ディスク(106-3(ID=B, C))が計算機-1(100-1)、計算機-2(100-2)、計算機-3(100-3)によって共有されるディスクである。

【0028】次に、表1に疎結合計算機システム構成例2に対応する構成定義情報の例を示す。構成定義情報は、ディスク要求処理部110に対して要求された処理がどの計算機のどのディスクに対応するか、処理の要求者が対応するディスクに対するアクセスの権利(アクセス権)を持っているか、セキュリティを向上させるためのパスワード等の構成定義を示す情報であり、テーブルとして構成される。仮想ディスクとは、プロセッサ101が発行する処理コマンドにおいて指定されるディスクの仮想的な名称である。この例では、疎結合計算機システム全体で8つのディスクがあり、VDISK-1~VDISK-8までの識別子を用いて指定される。

【0029】

【表1】

※すべての要求に対してパスワードチェックを行うことができる。

【0032】また、アクセス権の計算機からの要求であっても、特定のプロセス等からの要求に関してのみアクセスを許したい場合には、各要求毎にパスワードを付加し、チェックすることによりセキュリティを向上させることができる。

【0033】図3に疎結合計算機システム構成例2の計算機-1に対応する構成定義情報の例を示す。構成定義情報として、各計算機がどの実ディスク管理計算機に対応するかを示す情報を持つことで、計算機結合機構を介することなくアクセスできる仮想ディスクを識別することが可能となる。

【0034】表1及び図3の情報は、外部二次記憶装置に記録されていて初期化処理時に設定されることができる。また、ユーティリティツールにより、動作中に設定することもできる。

(7)

特開平9-6706

11

【0035】図4に、プロセッサ101から発行されるディスク処理要求コマンドの例を示す。

【0036】400は、ディスク処理要求コマンドである。仮想ディスクID401は、システム内のどのディスクを指定するかを示す識別子である。パスワード402は、図3にて示したパスワードチェックのために必要である。ディスク制御コマンド/データ403は、ディスク(106-1から106-N)を制御するためのコマンドとコマンドの実行に必要なデータであり、ディスク制御機構104を介して発行される。

【0037】図5に、ディスク処理終了時にプロセッサ101に返されるディスク処理ステータスの例を示す。500は、ディスク処理ステータスである。仮想ディスクID401は、図4の仮想ディスクID401と同一である。パスワード502は、図4のパスワード502と同じであるが、返送時には使用しなくてもよい。ディスクステータス/データは、ディスク(106-1から106-N)からプロセッサ101に返送されるコマンド終了時のステータス情報およびデータである。

【0038】図6に計算機結合機構アダプタパケットの例を示す。計算機結合機構では、パケットの配送を行う。送信側計算機ID601は送信側の計算機を示す識別子である。受信側計算機ID602は受信側の計算機を示す識別子である。データ603は転送されるデータである。パケットは、信頼性向上のために、パケットの前後に信頼性向上用冗長コードを付加することも好ましい。また、機密性を向上させるために、パケットを符号化することも好ましい。

【0039】図7に遠隔ディスク処理要求コマンドの例を示す。700は、遠隔ディスク処理要求コマンドである。要求元計算機ID701はディスクに対する処理を要求する計算機を示す識別子である。実ディスク管理計算機ID702は処理対象となるディスクを直接管理するディスク制御機構104を保持する計算機を示す識別子である。実ディスクID703は、処理対象となるディスクを示す識別子である。パスワード402及びディスク制御コマンド/データ403は、図4に示したものと同一である。

【0040】図8に遠隔ディスク処理ステータスの例を示す。800は、遠隔ディスク処理ステータスである。要求元計算機ID701はディスクに対する処理を要求した計算機を示す識別子であり、図7に示したものと同一である。実ディスク管理計算機ID702は処理対象となるディスクを直接管理するディスク制御機構104を保持する計算機を示す識別子であり、図7に示したものと同一である。実ディスクID703は、処理対象となるディスクを示す識別子であり、図7に示したものと同一である。パスワード502及びディスクステータス/データ503は、図5に示したものと同一である。

【0041】図9に実ディスク処理要求コマンドの例を

12

示す。900は、実ディスク処理要求コマンドである。実ディスクID703は、処理対象となるディスクを示す識別子であり、図7に示したものと同一である。ディスク制御コマンド/データ403は、ディスク(106-1から106-N)を制御するためのコマンドとコマンドの実行に必要なデータであり、図4に示したものと同一である。

【0042】図10に実ディスク処理ステータスの例を示す。1000は、実ディスク処理ステータスである。実ディスクID703は、処理対象となるディスクを示す識別子であり、図7に示したものと同一である。ディスクステータス/データ503は、ディスク(106-1から106-N)からプロセッサ101に返送されるコマンド終了時のステータス情報およびデータであり、図5に示したものと同一である。

【0043】図11にて、ディスク要求処理部動作のフローを説明する。まずステップ1110に示すようにプロセッサ101が発行したディスク処理要求コマンドを要求制御部111が受け付け、ステップ1120に進む。

【0044】次に、ステップ1120に示すように要求先判定部113が構成定義情報保持部114に格納されている構成定義情報(表1を参照)を参照して、仮想ディスクID401に対応する実ディスク管理計算機(表1を参照)を判定し、ステップ1130に進む。

【0045】次に、ステップ1130に示すように構成定義情報(表1及び図3を参照)を参照して仮想ディスクID401に対応する実ディスク管理計算機がディスク処理要求コマンドを発行したプロセッサ101と同一であるか否かを判定する。仮想ディスクID401に対応する実ディスク管理計算機がディスク処理要求コマンドを発行したプロセッサ101と同一でない場合には、ステップ1140に分岐し、仮想ディスクID401に対応する実ディスク管理計算機がディスク処理要求コマンドを発行したプロセッサ101と同一の場合には、ステップ1150に分岐する。

【0046】仮想ディスクID401に対応する実ディスク管理計算機がディスク処理要求コマンドを発行したプロセッサ101と同一でない場合には、ステップ1140に示すように遠隔要求処理動作を実行する。

【0047】図13と図14で遠隔要求処理動作について説明する。仮想ディスクID401に対応する実ディスク管理計算機がディスク処理要求コマンドを発行したプロセッサ101と同一の場合には、ステップ1150に示すように、要求許可判定部112がディスク処理要求コマンド400で示される処理対象となる仮想ディスクID401に対するアクセス権を構成定義情報中のアクセス権情報(表1を参照)を検索することでチェックし、ステップ1160に進む。図9Aに示した例では、アクセス権情報はアクセス権を持つ計算機を示してい

(8)

特開平9-6706

13

る。

【0048】次にステップ1160に示すようにアクセス権が無ければステップ1195に分岐し、アクセス権があればステップ1170に分岐する。

【0049】アクセス権がなければ、ステップ1195に示すようにその要求を拒否して終了する。要求を拒否して終了する場合には、要求制御部111がディスク要求処理ステータス500を作成してプロセッサ101に返送する。ディスク処理ステータス500は、仮想ディスクID401とパスワード402はディスク要求処理

コマンド400と同一であり、ディスクステータスとして、アクセス権がないことを通知するためのステータスコードを設定する。

【0050】アクセス権があれば、ステップ1170に示すように要求許否判定部112が、ディスク処理要求コマンド400のパスワード402と構成定義情報保持部114に保持された構成定義情報のパスワード305とを比較することでパスワードチェックを行い、ステップ1180に進む。次にステップ1180では、パスワードが一致しなければステップ1195に分岐し、パスワードが一致すればステップ1190に分岐する。この例では、アクセス権がない場合とパスワードが不一致な場合を分けてステータス通知を行っていないが、障害検出機能を向上するために、ステータス通知を分けることも好ましい。

【0051】パスワードが一致すれば、ステップ1190に示すように、非遠隔要求処理を実行し終了する。ディスク要求処理部動作は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。非遠隔要求処理動作については図12にて説明する。

【0052】図12にて、非遠隔要求処理動作のフローを説明する。まずステップ1210に示すように、要求制御部111がディスク処理要求コマンド400から、実ディスク処理要求コマンド900を作成する。構成定義情報(表1を参照)を参照して、ディスク処理要求コマンド400の仮想ディスクID401に対応する実ディスクID703に変換することで、ディスク処理要求コマンド400から実ディスク処理要求コマンド900を作成することができる。

【0053】次にステップ1220に示すように、要求制御部111がディスク制御機構共有制御部115に対して実ディスク処理要求コマンド900を発行する。

【0054】次にステップ1230に示すように、ディスク制御機構共有制御部、遠隔要求処理部121と要求制御部111とから発行された実ディスク処理要求コマンド900をシリアルライズし、ディスク制御機構104に対して発行する。実ディスク処理要求コマンド900をシリアルライズする処理は、図15にて説明する。

【0055】次にステップ1240に示すように、ディ

14

スク制御機構104が実ディスク処理要求コマンド900を用いてディスク(106-1~106-N)を制御する。

【0056】次にステップ1250に示すように、実ディスク処理要求コマンド900によるディスク(106-1~106-N)に対する制御の返り値として、ディスク制御機構104がディスク制御機構共有制御部115に実ディスク処理ステータス1000を返す。

【0057】次にステップ1260に示すように、ディスク制御機構共有制御部115が実ディスク処理ステータス1000を要求制御部111に返す。

【0058】次にステップ1270に示すように、要求制御部111は、構成定義情報(図3Aを参照)を参照して、実ディスク処理ステータス1000の実ディスクID703に対応する仮想ディスクID401を検索することで、ディスク処理ステータス500を作成する。

【0059】次にステップ1280に示すように、要求制御部111は、ディスク処理ステータス500をプロセッサ101に返送し、非遠隔要求処理を終了する。非遠隔要求処理動作は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。

【0060】図13にて、遠隔要求処理動作のフローを説明する。まず、ステップ1310に示すように遠隔要求発行部122がディスク処理要求コマンド400と構成定義情報とから、遠隔ディスク処理要求コマンド700を作成する。構成定義情報(表1を参照)を参照して、ディスク処理要求コマンド400の仮想ディスクID401に対応する実ディスクID703に変換することで、ディスク処理要求コマンド400から実ディスク処理要求コマンド900を作成することができる。

【0061】次にステップ1320に示すように遠隔要求発行部122が遠隔ディスク処理要求コマンド700の送信を計算機結合アダプタ制御部123に要求する。

【0062】次にステップ1330に示すように計算機結合機構アダプタ制御部123は、計算機結合機構アダプタ130を介して、実ディスク管理計算機(表1を参照)の遠隔要求処理部121に対して遠隔ディスク処理要求込コマンド700を転送する。

【0063】次にステップ1340に示すように遠隔要求処理部121は、要求許否判定部112に対して遠隔ディスク処理要求コマンド700のアクセス権のチェックを要求する。

【0064】次にステップ1350に示すように要求許否判定部112は、遠隔ディスク処理要求コマンド700を受け取り、遠隔ディスク処理要求コマンド700で示される処理対象となる実ディスク管理計算機ID702及び実ディスクID703に対するアクセス権を構成定義情報中のアクセス権情報(表1を参照)を検索することでチェックし、ステップ1360に進む。

(9)

特開平9-6706

15

【0065】次にステップ1360に示すように、アクセス権が無ければステップ1395に分岐し、アクセス権があればステップ1370に分岐する。

【0066】アクセス権がなければ、ステップ1395に示すようにその要求を拒否して終了する。

【0067】要求を拒否して終了する場合には、遠隔要求処理部121が遠隔ディスク処理ステータス800を作成し、遠隔要求発行部122に返送し、ディスク処理要求コマンドを発行したプロセッサ101にエラー通知を行う。この処理の詳細は、図16で説明する。

【0068】アクセス権があれば、ステップ1370に示すように要求許可判定部112が、遠隔ディスク処理要求コマンド700のパスワード402と構成定義情報保持部114に保持された構成定義情報のパスワード305とを比較することでパスワードチェックを行い、ステップ1380に進む。次にステップ1380では、パスワードが一致しなければステップ1395に分岐し、パスワードが一致すればステップ1390に分岐する。この例では、アクセス権がない場合とパスワードが不一致な場合を分けてステータス通知を行っていないが、障害検出機能を向上するために、ステータス通知を分けることも好ましい。

【0069】パスワードが一致すれば、ステップ1390に示すように、遠隔要求時実ディスク処理を実行し終了する。遠隔要求処理動作は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。遠隔要求時実ディスク処理については図14にて説明する。

【0070】図14にて遠隔要求時実ディスク処理のフローを説明する。まず、ステップ1410に示すように、遠隔要求処理部121が遠隔ディスク処理要求コマンド700から実ディスク処理要求コマンド900を作成し、ステップ1420に進む。実ディスク処理要求コマンドの実ディスクID703は、遠隔ディスク要求コマンド700の実ディスクID703と同一であり、実ディスク処理要求コマンドのディスク制御コマンド/データ403は、遠隔ディスク要求コマンド700のディスク制御コマンド/データ403と同一である。

【0071】次にステップ1420に示すように、遠隔要求処理部121がディスク制御機構共有制御部115に実ディスク処理要求コマンド900を発行し、ステップ1430に進む。

【0072】次にステップ1430に示すように、ディスク制御機構共有制御部115では、遠隔要求処理部121と要求制御部111とから発行された実ディスク処理コマンド900をシリアルライズして、ディスク制御機構104に対して発行し、ステップ1440に進む。ディスク制御機構共有制御部115の動作については、図15にて説明する。

【0073】次にステップ1440に示すようにディス

16

ク制御機構104が実ディスク処理要求コマンド900を用いてディスク(106-1~106-N)を制御し、ステップ1450に進む。

【0074】次にステップ1450に示すように、ディスク制御機構104は、実ディスク処理要求コマンド900によるディスク(106-1~106-N)の制御の返り値として、ディスク制御機構共有制御部115に実ディスク処理ステータス1000を返し、ステップ1460に進む。

10 【0075】次にステップ1460に示すように、ディスク制御機構共有部115が、実ディスク処理ステータス1000を遠隔要求処理部121に返し、ステップ1470に進む。

【0076】次にステップ1470に示すように、遠隔要求処理部121は、構成定義情報に基づき、実ディスク処理ステータス1000から、遠隔ディスク処理ステータス800を作成し、ステップ1480に進む。

20 【0077】次にステップ1480に示すように、遠隔要求処理部121が遠隔ディスク処理ステータス800の送信を計算機結合機構アダプタ123に要求し、ステップ1490に進む。

【0078】次にステップ1490に示すように、計算機結合機構アダプタ123は、計算機結合機構アダプタ130を介して、要求元計算機の遠隔要求発行部122に対して、遠隔ディスク処理ステータス800を転送し、ステップ1492に進む。

30 【0079】次にステップ1492に示すように、遠隔要求発行部122は、構成定義情報(図3Aを参照)を参照して、遠隔ディスク処理ステータス800の実ディスク管理計算機ID702及実ディスクID703対応する仮想ディスクID401を検索することでディスク処理ステータス500を作成し、ステップ1495に進む。

【0080】次にステップ1495に示すように、遠隔要求発行部122は、ディスク処理ステータス500をプロセッサ101に返送し、終了する。遠隔要求時実ディスク処理は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。

40 【0081】図15は、ディスク制御機構共有制御部115の実施例である。ディスク制御機構共有制御部115は、遠隔要求処理部対応要求保持部1510、要求制御部対応要求保持部1520、ディスク要求シリアルライズ部1530、実ディスク制御部1540から構成されている。

【0082】遠隔要求処理部対応要求保持部1510は、遠隔要求処理部121から発行された実ディスク要求コマンド900及び実ディスク処理ステータス1000を保持する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0083】要求制御部対応要求保持部1520は、要

(10)

特開平9-6706

17

求制御部111から発行された実ディスク要求コマンド900及び実ディスク処理ステータス1000を保持する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。

【0084】ディスク要求シリアルライズ部1530は、遠隔要求処理部121及び要求制御部111から発行された実ディスク要求コマンド900をシリアルライズして、実ディスク制御部1540に提供する部分であり、たとえば、TTL、CMOS等の論理によって構成できる。各要求のシリアルライズの例として、ラウンドロビン方式を採用し、遠隔要求処理部対応要求保持部1510にある要求と要求制御部対応要求保持部1520にある要求を交互に取り出す方法がある。

【0085】実ディスク制御部1540は、実ディスク要求コマンド900の発行及び実ディスク処理ステータス1000の受け取りを行う部分である。実ディスク要求コマンド900の発行は、ディスク要求シリアルライズ部1530によって出力された順番にしたがって、ディスク制御機構104に対して発行する。また、実ディスク処理ステータス1000の受け取りは、ディスク制御機構104が発行した順番にしたがって受け取る。受け取った実ディスク処理ステータス1000は、ディスク要求シリアルライズ部1530を介して、遠隔要求処理部対応要求保持部1510及び要求制御部対応要求保持部1520に転送される。遠隔要求処理部対応要求保持部1510からは、遠隔要求処理部121に実ディスク処理ステータス1000が返送される。要求制御部対応要求保持部1520からは、要求制御部111に実ディスク処理ステータス1000が返送される。ディスク制御機構共有制御部は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。

【0086】図16にて、遠隔要求拒否処理のフローを説明する。まず、ステップ1610に示すように、遠隔要求処理部121は、アクセス権エラー及びパスワードエラーに対するステータスと遠隔ディスク処理要求コマンド700とから、遠隔ディスク処理ステータス800を作成し、ステップ1620に進む。遠隔ディスク処理要求コマンド700と遠隔ディスク処理ステータス800は、ディスク制御コマンド/データ403とディスクステータス/データ503以外は、同一である。次にステップ1620に示すように、遠隔要求処理部121が遠隔ディスク処理ステータス800の送信を計算機結合機構アダプタ制御部123に要求し、ステップ1630に進む。次にステップ1630に示すように、計算機結合機構アダプタ123は、計算機結合機構アダプタ130を介して、要求元計算機の遠隔要求発行部122に対して、遠隔ディスク処理ステータス800を転送し、ステップ1640に進む。

【0087】次にステップ1640に示すように、遠隔

18

要求発行部122は、構成定義情報(図3Aを参照)を参照して、遠隔ディスク処理ステータス800の実ディスク管理計算機ID702及実ディスクID703対応する仮想ディスクID401を検索することでディスク処理ステータス500を作成し、ステップ1650に進む。

【0088】次にステップ1650に示すように、遠隔要求発行部122は、ディスク処理ステータス500をプロセッサ101に返送し、終了する。遠隔要求拒否処理は、ディスク要求処理部110にプロセッサを保持して、ソフトウェアによって実現されてもよい。

【0089】

【発明の効果】以上本発明によれば、共有対象のハードディスクの接続するバスとリモートアクセスのための接続機構の間の処理要求を必要に応じて転送することで、必要な場合にしか接続機構に対するアクセスが発生せず、要求の発生したプロセッサがハードディスクを直接制御できるため、高性能な共有ディスクを有する疎結合計算機システムを構築することが可能となる。

【図面の簡単な説明】

【図1】疎結合計算機システム構成の第1の例のブロック図。

【図2】疎結合計算機システム構成の第2の例のブロック図。

【図3】構成定義情報例概念図。

【図4】ディスク処理要求コマンド例概念図。

【図5】ディスク処理ステータス例概念図。

【図6】計算機結合機構アダプタバケット例概念図。

【図7】遠隔ディスク処理要求コマンド例概念図。

【図8】遠隔ディスク処理ステータス例概念図。

【図9】実ディスク処理要求コマンド例概念図。

【図10】実ディスク処理ステータス例概念図。

【図11】ディスク要求処理部動作フロー図。

【図12】非遠隔要求処理フロー図。

【図13】遠隔要求処理フロー図。

【図14】遠隔要求時実ディスク処理フロー図。

【図15】ディスク制御機構共有制御部のブロック図。

【図16】遠隔要求拒否処理フロー図。

【符号の説明】

(100-1から100-N) : 計算機

101 : プロセッサ、102 : メモリ、103 : I/O装置

104 : ディスク制御機構、105 : システムバス

(106-1から106-N) : ディスク

110 : ディスク要求処理部、111 : 要求制御部、1

12 : 要求拒否判定部

113 : 要求先判定部、114 : 構成定義情報保持部

115 : ディスク制御機構共有制御部、116 : 内部バス

120 : 遠隔ディスク制御部、121 : 遠隔要求処理

(11)

特開平9-6706

19

部、122:遠隔要求発行部

123:計算機結合機構アダプタ制御部、124:内部バス

140:計算機結合機構

301:仮想ディスク、302:実ディスク管理計算機

303:管理計算機内対応ディスク、304:アクセス権情報、305:パスワード

400:ディスク処理要求コマンド、401:仮想ディスクID

402:パスワード、403:ディスク制御コマンド / 10データ

500:ディスク処理ステータス、502:パスワード

601:送信側計算機ID、602:受信側計算機ID、603:データ

700:遠隔ディスク処理要求コマンド、701:要求

20

元計算機ID

702:実ディスク管理計算機ID、703:実ディスクID

800:遠隔ディスク処理ステータス

900:実ディスク処理要求コマンド

1000は:実ディスク処理ステータス

1110~1195:処理ステップ

1210~1280:処理ステップ

1310~1395:処理ステップ

1410~1495:処理ステップ

1510:遠隔要求処理部対応要求保持部、1520:要求制御部対応要求保持部

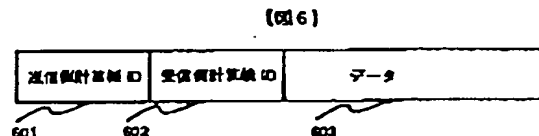
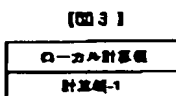
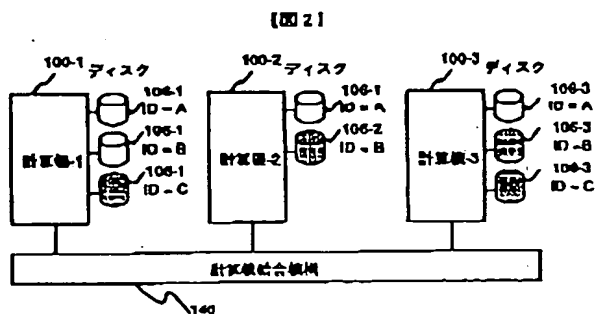
1530:ディスク要求シリアルライズ部、1540:実ディスク制御部

1610~1650:処理ステップ

【図2】

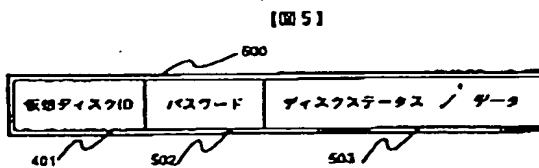
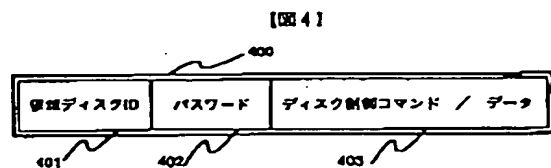
【図3】

【図6】



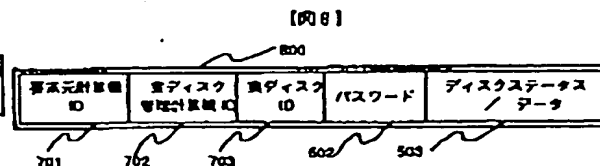
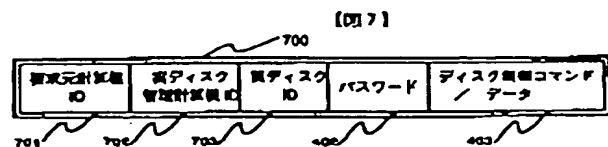
【図4】

【図5】



【図7】

【図8】

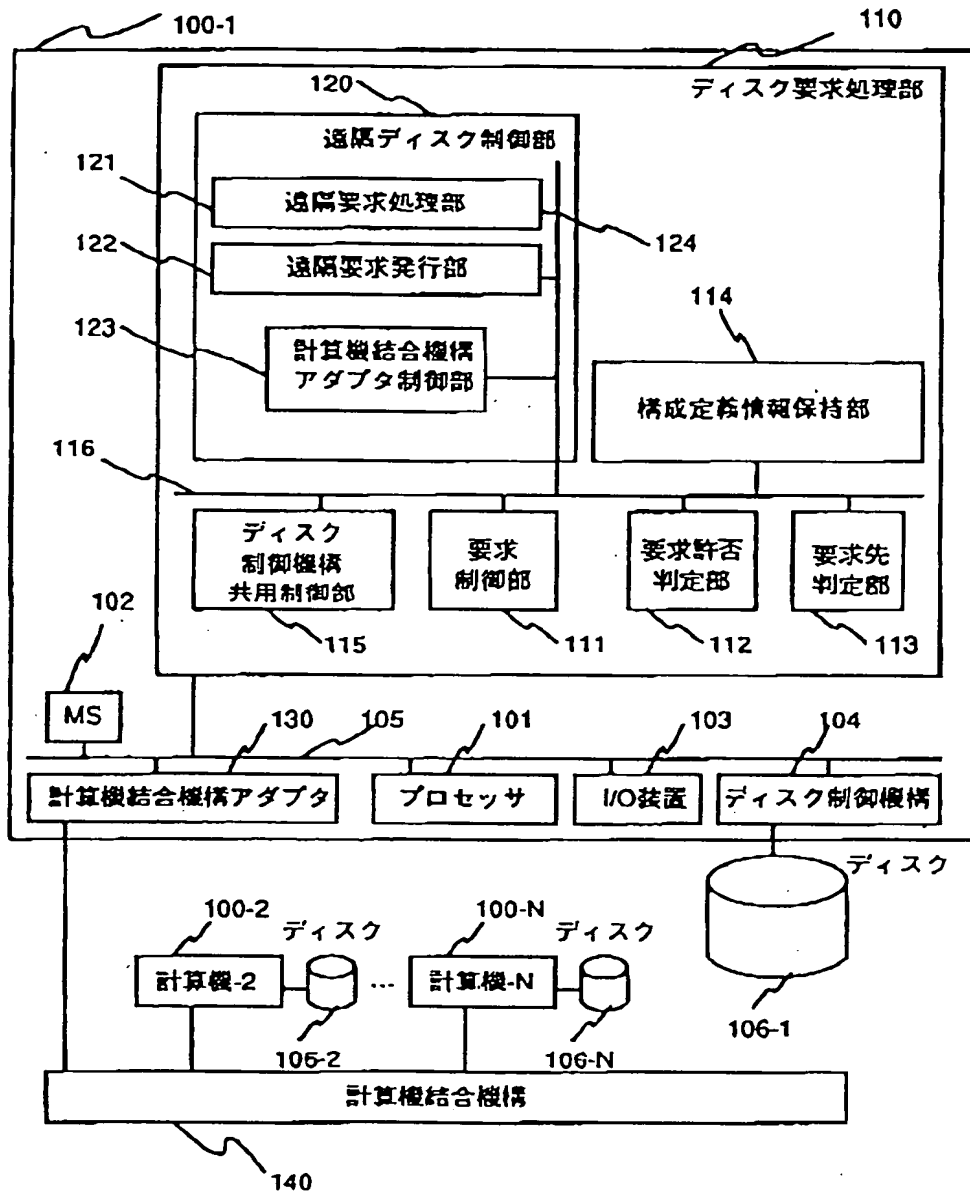


(12)

特開平9-6706

【図1】

【図1】

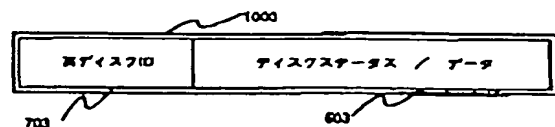
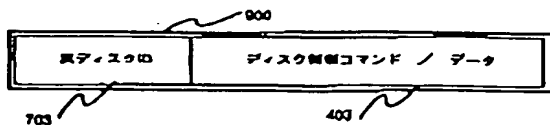


【図9】

【図10】

【図9】

【図10】

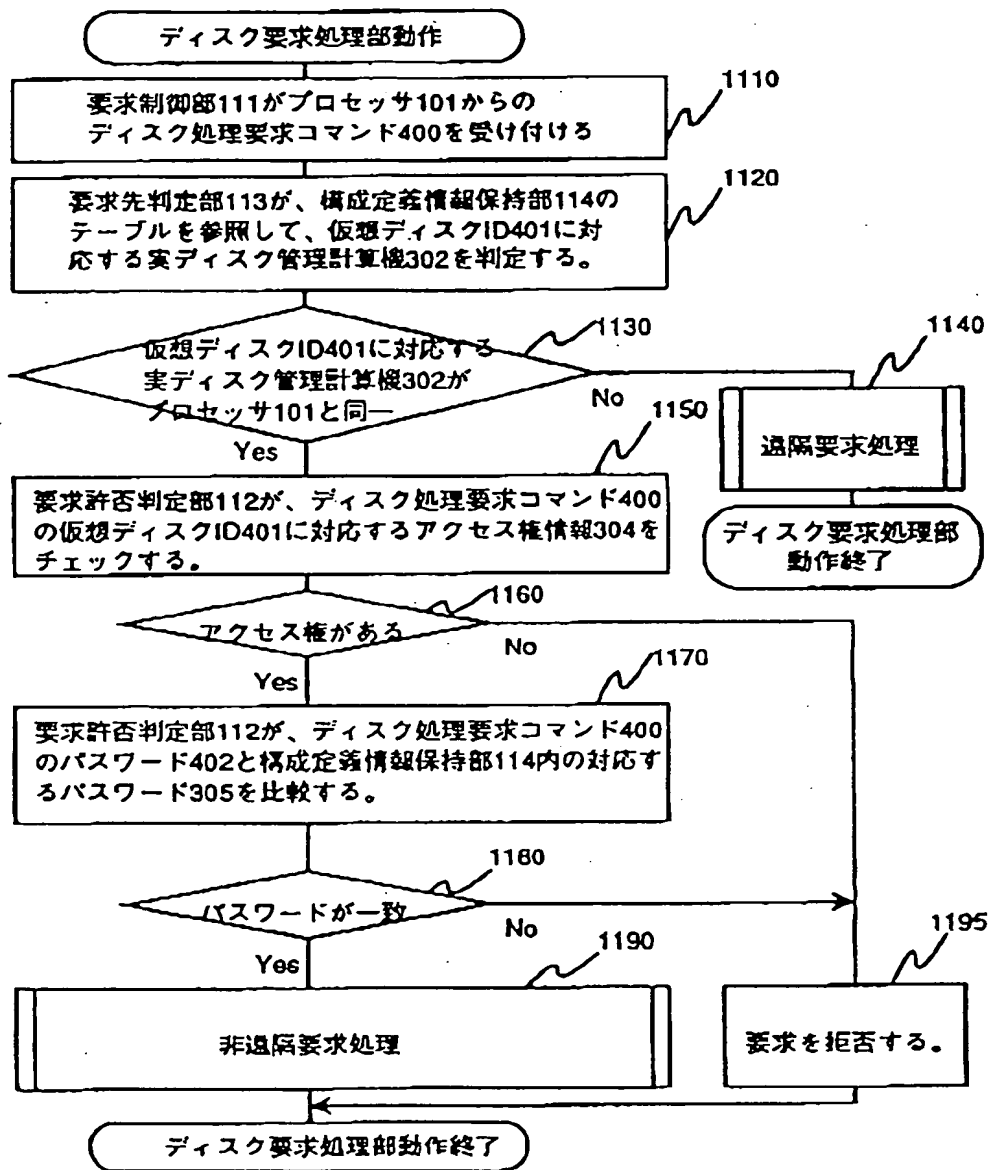


(13)

特開平9-6706

【図11】

【図11】

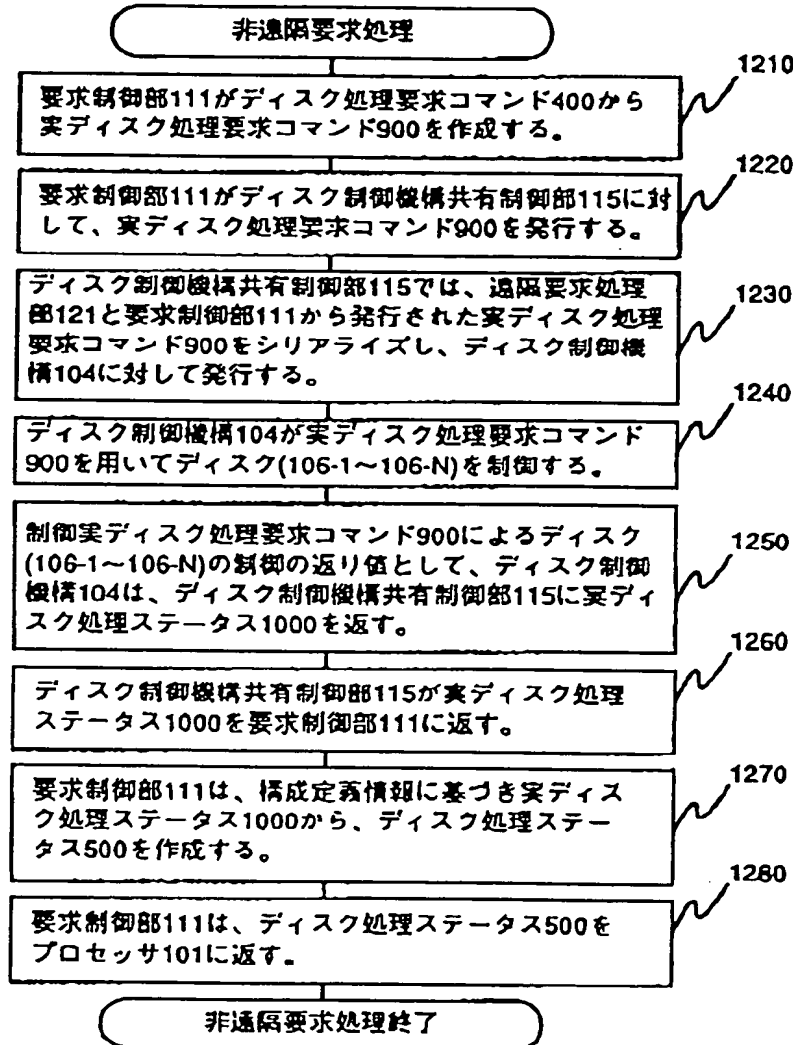


(14)

特開平9-6706

【図12】

【図12】

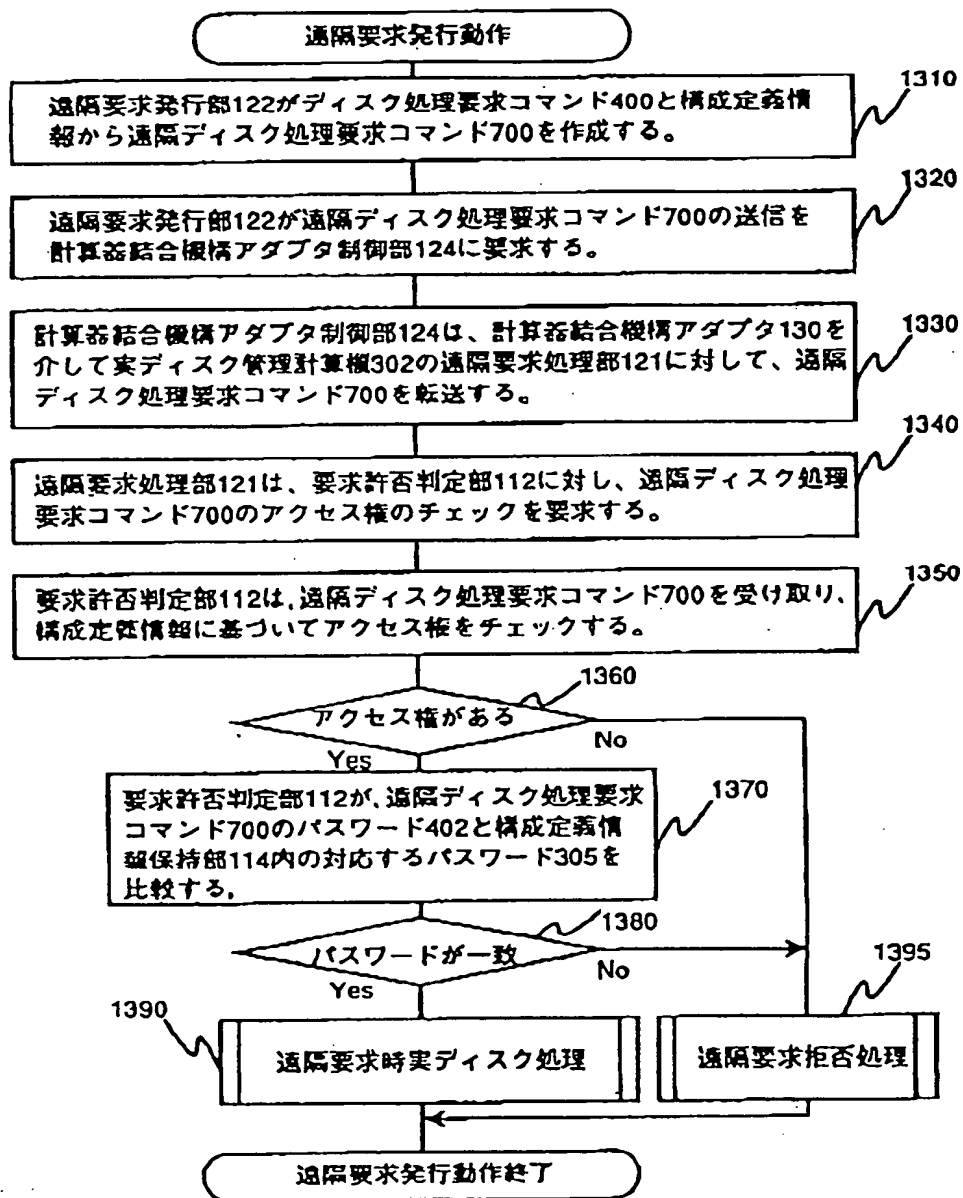


(15)

特開平9-6706

【図13】

【図13】

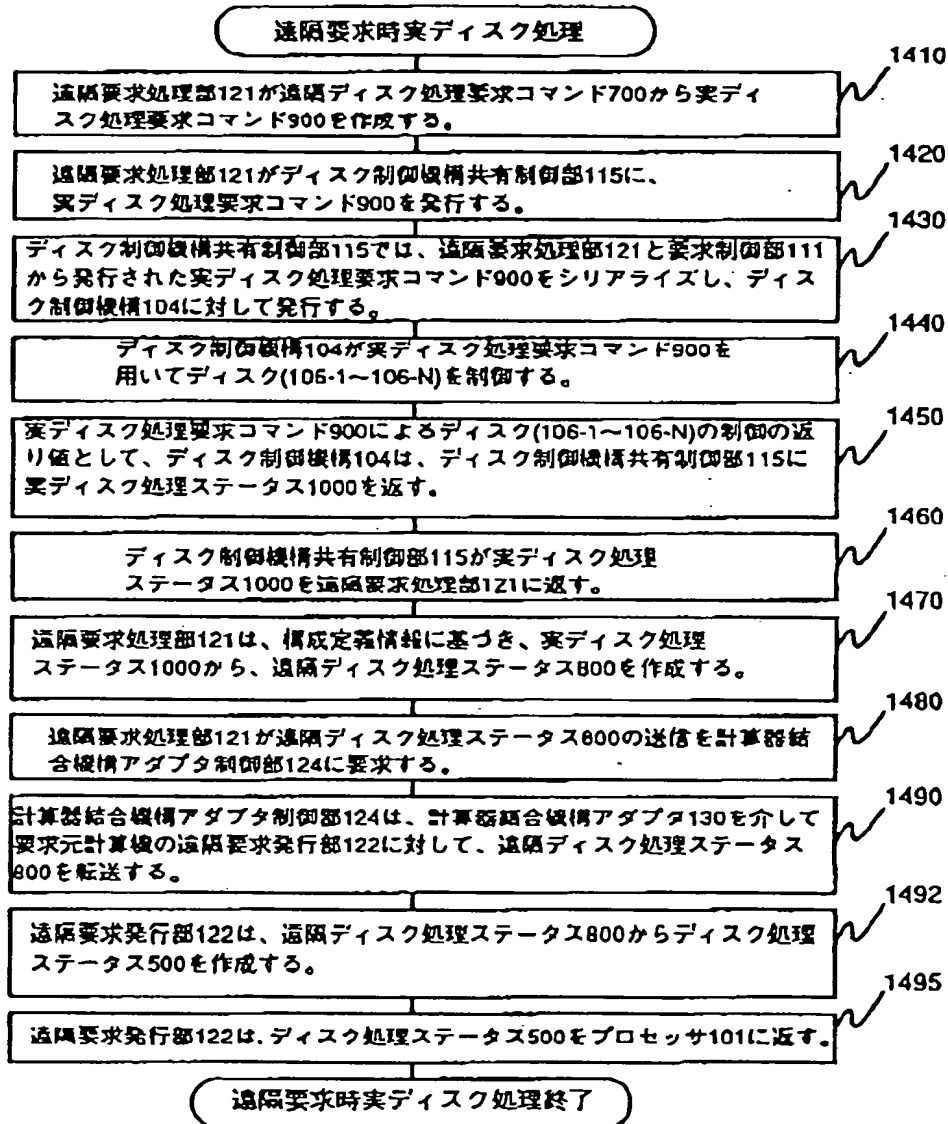


(16)

特開平9-6706

【図14】

【図14】

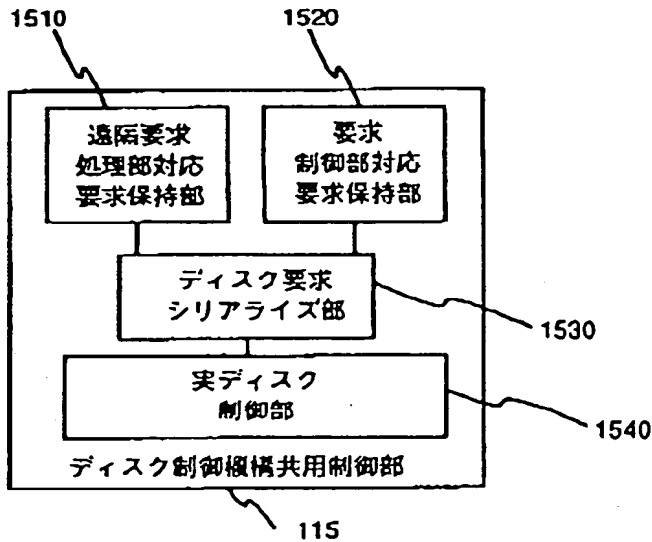


(17)

特開平9-6706

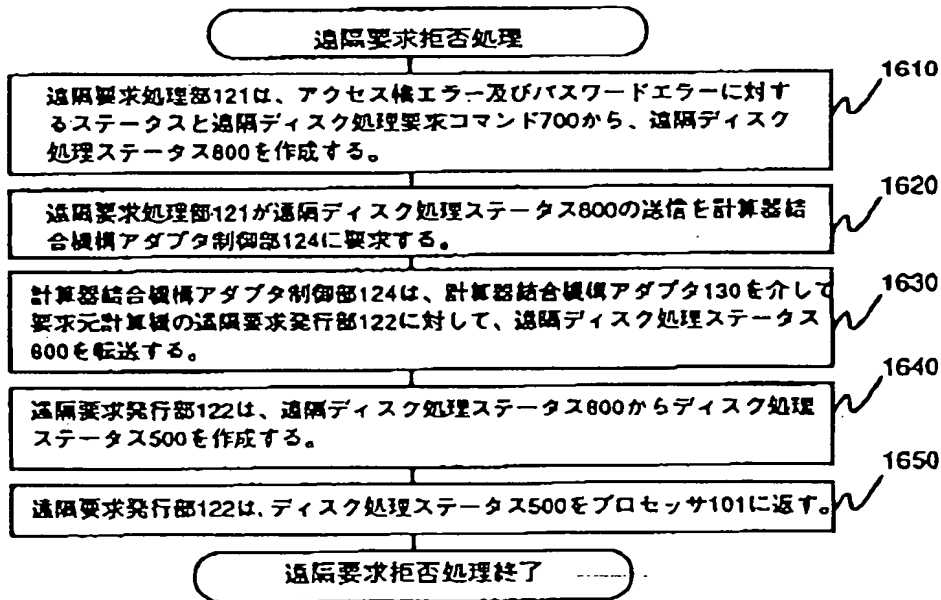
【図15】

【図15】



【図16】

【図16】



フロントページの続き

(72)発明者 堀川 和雄
東京都国分寺市東恋ヶ程1丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 林 剛久
東京都国分寺市東恋ヶ程1丁目280番地
株式会社日立製作所中央研究所内

(18)

特開平9-6706

(72)発明者 岩本 博志
神奈川県海老名市下今泉810番地 株式会
社日立製作所オフィスシステム事業部内

(72)発明者 山田 公稔
神奈川県横浜市戸塚区戸塚町5030番地 株
式会社日立製作所ソフトウェア開発本部内